確率・統計基礎: 複数の確率変数、多変量正規分布

森 立平

複数の確率変数の確率密度関数

多変数の確率密度関数

$p_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}$ を n 次元確率変数 $\mathbf{X} = [X_1, \ldots, X_n]$ の確率密度関数とすると、任意の $A \subseteq \mathbb{R}^n$ について

$$\Pr(\mathbf{X} \in A) = \int_A p_{\mathbf{X}}(x) \mathrm{d}x.$$

独立な場合

確率変数 $X_1, \ldots, X_n$ が独立な場合、

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = p_{X_1}(x_1) p_{X_2}(x_2) \cdots p_{X_n}(x_n).$$

Example

$X_1 \sim U(0,1)$ で ($[0,1]$ 上の一様分布)、$X_2 = X_1$ のとき $[X_1, X_2]$ の確率密度関数は存在しない。$p_{X_1,X_2}(x_1, x_2)$ は $x_1 \neq x_2$ のとき 0 としてよいが、$p_{X_1,X_2}(x, x)$ をどのように定義しても $\int_{\mathbb{R}^2} p_{X_1,X_2}(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2 = 0$ になる。

周辺化

一つの確率変数 X_1 の確率密度関数は

$$\begin{aligned}\Pr(X_1 \in A) &= \Pr(X_1 \in A, X_2 \in \mathbb{R}, \dots, X_n \in \mathbb{R}) \\ &= \int_{\substack{x_1 \in A \\ x_2,\dots,x_n \in \mathbb{R}}} p_{X_1,\dots,X_n}(x_1, x_2, \dots, x_n)\mathrm{d}x \\ &= \int_{x_1 \in A} \left(\int_{x_2,\dots,x_n \in \mathbb{R}} p_{X_1,\dots,X_n}(x_1, x_2, \dots, x_n)\mathrm{d}x_2 \cdots \mathrm{d}x_n \right) \mathrm{d}x_1.\end{aligned}$$

よって X_1 の確率密度関数は

$$p_{X_1}(x_1) = \int_{x_2,\dots,x_n \in \mathbb{R}} p_{X_1,\dots,X_n}(x_1, x_2, \dots, x_n)\mathrm{d}x_2 \cdots \mathrm{d}x_n$$

期待値

確率変数のベクトルや行列に対する $\mathbb{E}[\cdot]$ は成分毎に期待値を取ることを表す。

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \text{とすると} \qquad \mathbb{E}[\mathbf{X}] = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_n] \end{bmatrix}$$

である。

分散共分散行列

Definition (分散共分散行列)

確率変数 $X_1, \ldots, X_n$ の分散共分散行列 (もしくは共分散行列) $\mathrm{Var}[\mathbf{X}] \in \mathbb{R}^{n\times n}$ は $(i,\ j)$ 成分に $\mathrm{Cov}\left[X_i,\ X_j\right] = \mathbb{E}\left[X_i X_j\right] - \mathbb{E}[X_i]\,\mathbb{E}\left[X_j\right]$ を持つ行列である。

$$\mathrm{Var}[\mathbf{X}] = \begin{bmatrix} \mathbb{V}[X_1] & \mathrm{Cov}[X_1, X_2] & \mathrm{Cov}[X_1, X_3] \\ \mathrm{Cov}[X_2, X_1] & \mathbb{V}[X_2] & \mathrm{Cov}[X_2, X_3] \\ \mathrm{Cov}[X_3, X_1] & \mathrm{Cov}[X_3, X_2] & \mathbb{V}[X_3] \end{bmatrix}$$

分散共分散行列は以下のように表せる。

$$\mathbf{\Sigma} = \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] - \mathbb{E}[\mathbf{X}]\,\mathbb{E}[\mathbf{X}]^T$$

分散共分散行列の性質

- $\mathrm{Var}\left[\mathbf{X}\right]$ は対称行列である。
- $\mathrm{Var}\left[\mathbf{X}\right]$ は半正定値である。
- $\mathrm{Var}\left[\mathbf{X}+b\right] = \mathrm{Var}\left[\mathbf{X}\right] \quad \forall b \in \mathbb{R}^n.$
- $\mathrm{Var}\left[A\mathbf{X}\right] = A\mathrm{Var}\left[\mathbf{X}\right]A^T \quad \forall A \in \mathbb{R}^{m\times n}.$

線形変換の分散共分散行列

行列 $A \in \mathbb{R}^{m \times n}$ と $n \times k$ の確率変数行列 $\mathbf{R}$ について $\mathbb{E}[A\mathbf{R}] = A\mathbb{E}[\mathbf{R}]$ である。よって、N 次元確率変数ベクトル $\mathbf{X}$ について

$$\begin{aligned}
&\mathbb{E}\left[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T\right] \\
=&\mathbb{E}\left[\mathbf{X}\mathbf{X}^T - \mathbf{X}\mathbb{E}[\mathbf{X}]^T - \mathbb{E}[\mathbf{X}]\mathbf{X}^T + \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T\right] \\
=&\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T = \mathrm{Var}[\mathbf{X}]
\end{aligned}$$

よって $\mathrm{Var}[\mathbf{X} + b] = \mathrm{Var}[\mathbf{X}]$ である。

$$\begin{aligned}
\mathrm{Var}[A\mathbf{X}] &= \mathbb{E}\left[A\mathbf{X}(A\mathbf{X})^T\right] - \mathbb{E}[A\mathbf{X}]\mathbb{E}[A\mathbf{X}]^T \\
&= A\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]A^T - A\mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T A^T \\
&= A\mathrm{Var}[\mathbf{X}]A^T
\end{aligned}$$

分散共分散行列の半正定値性 I

$$
\begin{bmatrix}
1 & \mathbb{E}[X_1] & \mathbb{E}[X_2] & \mathbb{E}[X_3] \\
\mathbb{E}[X_1] & \mathbb{E}\left[X_1^2\right] & \mathbb{E}[X_1X_2] & \mathbb{E}[X_1X_3] \\
\mathbb{E}[X_2] & \mathbb{E}[X_2X_1] & \mathbb{E}\left[X_2^2\right] & \mathbb{E}[X_2X_3] \\
\mathbb{E}[X_3] & \mathbb{E}[X_3X_1] & \mathbb{E}[X_3X_2] & \mathbb{E}\left[X_3^2\right]
\end{bmatrix}
$$

$$
=\mathbb{E}\left[\begin{bmatrix}
1 & X_1 & X_2 & X_3 \\
X_1 & X_1^2 & X_1X_2 & X_1X_3 \\
X_2 & X_2X_1 & X_2^2 & X_2X_3 \\
X_3 & X_3X_1 & X_3X_2 & X_3^2
\end{bmatrix}\right]
$$

$$
=\mathbb{E}\left[\begin{bmatrix} 1 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}\begin{bmatrix} 1 & X_1 & X_2 & X_3 \end{bmatrix}\right] \geq 0 \qquad \text{(半正定値行列の期待値は半正定値)}
$$

任意の $a \in \mathbb{R}^n$ と $n \times n$ 半正定値ランダム行列 R について
$a^T\mathbb{E}[R]\,a = \mathbb{E}\left[a^TRa\right] \geq 0$ なので $\mathbb{E}[R] \geq 0$.

分散共分散行列の半正定値性 II

$$\begin{bmatrix} -\mathbb{E}[X_1] & 1 & 0 & 0 \\ -\mathbb{E}[X_2] & 0 & 1 & 0 \\ -\mathbb{E}[X_3] & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \mathbb{E}[X_1] & \mathbb{E}[X_2] & \mathbb{E}[X_3] \\ \mathbb{E}[X_1] & \mathbb{E}[X_1^2] & \mathbb{E}[X_1X_2] & \mathbb{E}[X_1X_3] \\ \mathbb{E}[X_2] & \mathbb{E}[X_2X_1] & \mathbb{E}[X_2^2] & \mathbb{E}[X_2X_3] \\ \mathbb{E}[X_3] & \mathbb{E}[X_3X_1] & \mathbb{E}[X_3X_2] & \mathbb{E}[X_3^2] \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 & \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2] & \mathbb{E}[X_1X_3] - \mathbb{E}[X_1]\mathbb{E}[X_3] \\ 0 & \mathbb{E}[X_2X_1] - \mathbb{E}[X_2]\mathbb{E}[X_1] & \mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2 & \mathbb{E}[X_2X_3] - \mathbb{E}[X_2]\mathbb{E}[X_3] \\ 0 & \mathbb{E}[X_3X_1] - \mathbb{E}[X_3]\mathbb{E}[X_1] & \mathbb{E}[X_3X_2] - \mathbb{E}[X_3]\mathbb{E}[X_2] & \mathbb{E}[X_3^2] - \mathbb{E}[X_3]^2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & \mathbb{V}[X_1] & \mathrm{Cov}[X_1, X_2] & \mathrm{Cov}[X_1, X_3] \\ 0 & \mathrm{Cov}[X_2, X_1] & \mathbb{V}[X_2] & \mathrm{Cov}[X_2, X_3] \\ 0 & \mathrm{Cov}[X_3, X_1] & \mathrm{Cov}[X_3, X_2] & \mathbb{V}[X_3] \end{bmatrix} \begin{bmatrix} -\mathbb{E}[X_1] & -\mathbb{E}[X_2] & -\mathbb{E}[X_3] \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \mathrm{Var}[\mathbf{X}]$$

$A \geq 0$ のとき任意の行列 B について $BAB^T \geq 0$ なので $\mathrm{Var}[\mathbf{X}] \geq 0$ である。

多変量正規分布

Definition (多変量正規分布)

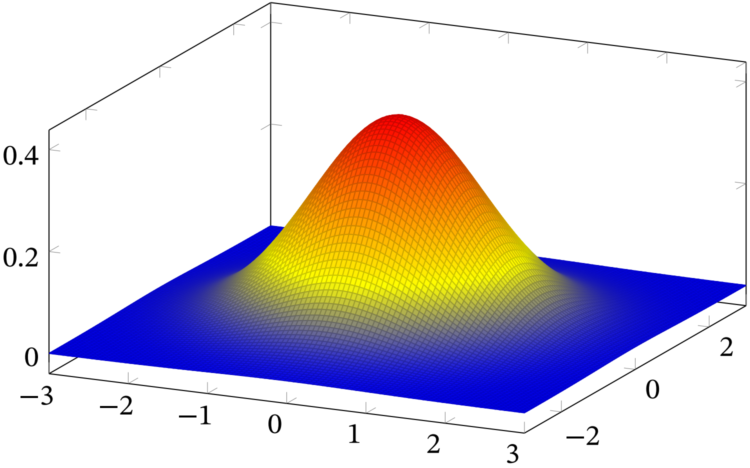
平均 $\mu \in \mathbb{R}^n$、分散共分散行列 $\Sigma \in \mathbb{R}^{n\times n}_{>0}$ の多変量正規分布 $N(\mu, \Sigma)$ に従う n 次元確率変数 $\mathbf{Z}$ の確率密度関数は

$$p_{\mathbf{Z}}(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}(x-\boldsymbol{\mu})^T \Sigma^{-1}(x-\boldsymbol{\mu})} \qquad \text{for } x \in \mathbb{R}^n$$

で定義される。
また、$n \times n$ 単位行列 I_n について $N(0, I_n)$ は n 次元標準正規分布と呼ばれる。

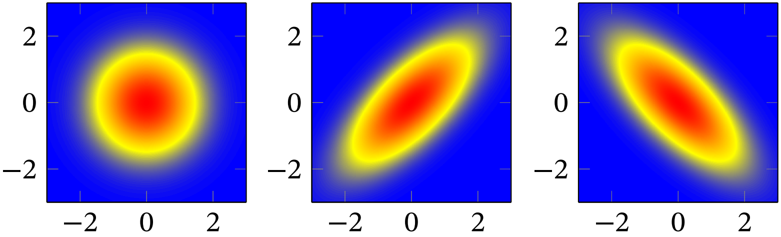
確率密度関数

$$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



確率密度関数

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sqrt{2} & -1 \\ -1 & \sqrt{2} \end{bmatrix}$$



密度関数の積分

多変量正規分布の密度関数を積分して 1 になることを確認する。Σ は対称行列なのである直交行列 V と対角行列 D が存在して $\Sigma = V^T DV$ と表せる (スペクトル分解)。また Σ は正定値なので、D の対角成分は正である。

$$
\begin{aligned}
&\int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} \mathrm{d}x \\
&= \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}x^T\Sigma^{-1}x} \mathrm{d}x \\
&= \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det(D)}} \mathrm{e}^{-\frac{1}{2}(Vx)^T D^{-1}(Vx)} \mathrm{d}x \qquad (\Sigma = V^T DV) \\
&= \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det(D)}} \mathrm{e}^{-\frac{1}{2}y^T D^{-1}y} \frac{1}{|\det(V)|} \mathrm{d}y \qquad (y = Vx) \\
&= \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det(D)}} \mathrm{e}^{-\frac{1}{2}z^T z} \frac{1}{\left|\det(\sqrt{D^{-1}})\right|} \mathrm{d}z \qquad (z = \sqrt{D^{-1}}y) \\
&= \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n}} \mathrm{e}^{-\frac{1}{2}z^T z} \mathrm{d}z = \prod_{k=1}^{n} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z_k^2}{2}} \mathrm{d}z_k \right) = 1
\end{aligned}
$$

多変数のモーメント母関数、キュムラント母関数、特性関数

Definition (多変数モーメント母関数)

n 次元確率変数 $\mathbf{X}$ のモーメント母関数 $\varphi_{\mathbf{X}} \colon \mathbb{R}^n \to \mathbb{R}$ を以下で定義する。

$$M_{\mathbf{X}}(t) := \mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{X}\rangle}\right].$$

ここで $\langle t, \mathbf{X}\rangle := \sum_{k=1}^{n} t_k X_k$ である。
またキュムラント母関数を $K_{\mathbf{X}}(t) := \log M_{\mathbf{X}}(t)$ と定義する。

Definition (多変数特性関数)

n 次元確率変数 $\mathbf{X}$ の特性関数 $\varphi_{\mathbf{X}} \colon \mathbb{R}^n \to \mathbb{C}$ を以下で定義する。

$$\varphi_{\mathbf{X}}(t) := \mathbb{E}\left[\mathrm{e}^{i\langle t, \mathbf{X}\rangle}\right].$$

多変数のモーメント母関数、キュムラント母関数、特性関数の性質

- $M_{\mathbf{X}}(t)$ が原点を含む開集合で存在するとき、この範囲で無限回微分できる (ルベーグ積分に基づいた議論が必要)。このとき、t に関する微分と期待値の順番を交換できる。

$$
\begin{aligned}
\left.\frac{\partial M_{\mathbf{X}}(t)}{\partial t_k}\right|_{t=0} &= \mathbb{E}\left[X_k e^{\langle t, \mathbf{X}\rangle}\right]\Big|_{t=0} = \mathbb{E}\left[X_k\right] \\
\left.\frac{\partial^2 M_{\mathbf{X}}(t)}{\partial t_k \partial t_\ell}\right|_{t=0} &= \mathbb{E}\left[X_k X_\ell e^{\langle t, \mathbf{X}\rangle}\right]\Big|_{t=0} = \mathbb{E}\left[X_k X_\ell\right] \\
\left.\frac{\partial K_{\mathbf{X}}(t)}{\partial t_k}\right|_{t=0} &= \left.\frac{\mathbb{E}\left[X_k e^{\langle t, \mathbf{X}\rangle}\right]}{\mathbb{E}\left[e^{\langle t, \mathbf{X}\rangle}\right]}\right|_{t=0} = \mathbb{E}\left[X_k\right] \\
\left.\frac{\partial^2 K_{\mathbf{X}}(t)}{\partial t_k \partial t_\ell}\right|_{t=0} &= \left.\frac{\mathbb{E}\left[X_k X_\ell e^{\langle t, \mathbf{X}\rangle}\right]}{\mathbb{E}\left[e^{\langle t, \mathbf{X}\rangle}\right]}\right|_{t=0} - \left(\left.\frac{\mathbb{E}\left[X_k e^{\langle t, \mathbf{X}\rangle}\right]}{\mathbb{E}\left[e^{\langle t, \mathbf{X}\rangle}\right]}\right|_{t=0}\right)\left(\left.\frac{\mathbb{E}\left[X_\ell e^{\langle t, \mathbf{X}\rangle}\right]}{\mathbb{E}\left[e^{\langle t, \mathbf{X}\rangle}\right]}\right|_{t=0}\right) \\
&= \mathbb{E}\left[X_k X_\ell\right] - \mathbb{E}\left[X_k\right]\mathbb{E}\left[X_\ell\right] = \mathrm{Cov}\left[X_k, X_\ell\right].
\end{aligned}
$$

- $\mathbf{X}$ と $\mathbf{Y}$ の分布が同じ $\Longleftrightarrow$ $\varphi_{\mathbf{X}} = \varphi_{\mathbf{Y}}$ (ルベーグ積分に基づいた議論が必要)。

多変数のモーメント母関数の性質

- 任意の n 次元確率変数 $\mathbf{X}$ と $b \in \mathbb{R}^n$ について
$$\begin{aligned} M_{\mathbf{X}+b}(t) &= \mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{X}+b\rangle}\right] = \mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{X}\rangle}\mathrm{e}^{\langle t, b\rangle}\right] = \mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{X}\rangle}\right]\mathrm{e}^{\langle t, b\rangle} \\ &= M_{\mathbf{X}}(t)\mathrm{e}^{\langle t, b\rangle}. \end{aligned}$$

- 任意の n 次元確率変数 $\mathbf{X}$ と $A \in \mathbb{R}^{m\times n}$ について
$$M_{A\mathbf{X}}(t) = \mathbb{E}\left[\mathrm{e}^{\langle t, A\mathbf{X}\rangle}\right] = \mathbb{E}\left[\mathrm{e}^{\langle A^T t, \mathbf{X}\rangle}\right] = M_{\mathbf{X}}(A^T t).$$

- 任意の独立 n 次元確率変数 $\mathbf{X}$ と $\mathbf{Y}$ について
$$M_{\mathbf{X}+\mathbf{Y}}(t) = \mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{X}+\mathbf{Y}\rangle}\right] = \mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{X}\rangle}\right]\mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{Y}\rangle}\right] = M_{\mathbf{X}}(t)M_{\mathbf{Y}}(t).$$

多変量正規分布のキュムラント母関数

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{\langle t, \mathbf{Z}\rangle}\right] \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \mathrm{e}^{\langle t, x\rangle} \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}\left\langle \Sigma^{-\frac{1}{2}}(x-\mu), \Sigma^{-\frac{1}{2}}(x-\mu)\right\rangle + \left\langle \Sigma^{\frac{1}{2}} t, \Sigma^{-\frac{1}{2}} x\right\rangle} \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}\left\langle \Sigma^{-\frac{1}{2}}(x-\mu-\Sigma t), \Sigma^{-\frac{1}{2}}(x-\mu-\Sigma t)\right\rangle + \langle \mu, t\rangle + \frac{1}{2}\left\langle \Sigma^{\frac{1}{2}} t, \Sigma^{\frac{1}{2}} t\right\rangle} \mathrm{d}x \\
&= \mathrm{e}^{\langle \mu, t\rangle + \frac{1}{2}\langle \Sigma^{\frac{1}{2}} t, \Sigma^{\frac{1}{2}} t\rangle}.
\end{aligned}
$$

よって、$K_{\mathbf{Z}}(t) = \langle \mu,\, t\rangle + \frac{1}{2}\left\langle \Sigma^{\frac{1}{2}} t,\, \Sigma^{\frac{1}{2}} t\right\rangle$.

$$
\left.\frac{\partial K_{\mathbf{Z}}(t)}{\partial t_k}\right|_{t=0} = \mu_k
$$

$$
\left.\frac{\partial^2 K_{\mathbf{Z}}(t)}{\partial t_k \partial t_\ell}\right|_{t=0} = \Sigma_{k,\ell}.
$$

多変量正規分布の周辺化も多変量正規分布

Lemma

n 次元の確率変数 $\mathbf{Z} = [Z_1, \dots, Z_n] \sim N(\mu, \Sigma)$ について、
$\mathbf{Z}' := [Z_1, \dots, Z_{n-1}] \sim N(\mu', \Sigma')$. ここで、$\mu' \in \mathbb{R}^{N-1}$ は μ の第 n 成分を除いたもので Σ' は Σ の n 行目と n 列目を除いたもの。

Proof.

$$\begin{aligned}
\varphi_{\mathbf{Z}'}(t') &= \mathbb{E}\left[\mathrm{e}^{i\langle t', \mathbf{Z}' \rangle}\right] \\
&= \mathbb{E}\left[\mathrm{e}^{i\langle t', \mathbf{Z}' \rangle + 0 \cdot Z_n}\right] \\
&= \varphi_{\mathbf{Z}}(t', 0) = \mathrm{e}^{i\langle t', \mu' \rangle - \langle \sqrt{\Sigma'} t', \sqrt{\Sigma'} t' \rangle}.
\end{aligned}$$

□

多変量正規分布の一次変換

Lemma

n 次元の確率変数 $\mathbf{Z} \sim N(\mu, \Sigma)$ とランク m 行列 $A \in \mathbb{R}^{m \times n}$ と $b \in \mathbb{R}^m$ について、$A\mathbf{Z} + b \sim N(A\mu + b, A\Sigma A^T)$.

Proof.

$\varphi_{\mathbf{Z}}(t) = \mathrm{e}^{i\langle \mu, t\rangle - \frac{1}{2}\langle \Sigma^{\frac{1}{2}} t, \Sigma^{\frac{1}{2}} t\rangle}$ より、

$$
\begin{aligned}
\varphi_{A\mathbf{Z}+b}(t) &= \varphi_{A\mathbf{Z}}(t)\mathrm{e}^{i\langle t, b\rangle} = \varphi_{\mathbf{Z}}(A^T t)\mathrm{e}^{i\langle t, b\rangle} \\
&= \mathrm{e}^{i\langle \mu, A^T t\rangle - \frac{1}{2}\langle \Sigma^{\frac{1}{2}} A^T t, \Sigma^{\frac{1}{2}} A^T t\rangle + i\langle t, b\rangle} \\
&= \mathrm{e}^{i\langle A\mu, t\rangle - \frac{1}{2}\langle \Sigma^{\frac{1}{2}} A^T t, \Sigma^{\frac{1}{2}} A^T t\rangle + i\langle t, b\rangle} \\
&= \mathrm{e}^{i\langle A\mu + b, t\rangle - \frac{1}{2}\langle \Sigma^{\frac{1}{2}} A^T t, \Sigma^{\frac{1}{2}} A^T t\rangle} \\
&= \mathrm{e}^{i\langle A\mu + b, t\rangle - \frac{1}{2}\langle t, A\Sigma A^T t\rangle}.
\end{aligned}
$$

これは $N(A\mu + b,\, A\Sigma A^T)$ の特性関数である。 $\square$

逆関数法: 一様分布から一般の分布の作り方

Theorem (逆関数法)

$U \sim U(0,1)$ とすると、任意の確率変数 X と $F_X^{-1}(U)$ は同じ分布を持つ。ここで、F_X は X の累積分布関数であり

$$F_X^{-1}(p) := \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\} \qquad \forall p \in [0,1].$$

Proof.

$$\Pr(F_X^{-1}(U) \leq x) = \Pr(U \leq F_X(x)) = F_X(x)$$

最初の等号は以下の関係から従う。

$$F_X^{-1}(u) \leq x \iff u \leq F_X(x)$$

$\Longrightarrow$: F_X が右連続であることに注意すると $u \leq F_X(F_X^{-1}(u))$ なので

$$F_X^{-1}(u) \leq x \implies F_X(F_X^{-1}(u)) \leq F_X(x). \implies u \leq F_X(x)$$

$\Longleftarrow$: $F_X^{-1}(F_X(x)) \leq x$ より

$$u \leq F_X(x) \implies F_X^{-1}(u) \leq F_X^{-1}(F_X(x)) \implies F_X^{-1}(u) \leq x.$$

□

二次元標準正規分布の極座標表示

$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ について極座標表示 $[R, \Theta]$ を考える。つまり、

$$R := \sqrt{X^2 + Y^2} \qquad \Theta := \arg(X, Y)$$
$$X = R\cos\Theta \qquad Y = R\sin\Theta$$

$$\begin{aligned}
&\Pr([R, \Theta] \in A) \\
&\quad= \int_{[r(x,y),\,\theta(x,y)]\in A} \frac{1}{2\pi} \mathrm{e}^{-\frac{1}{2}(x^2+y^2)} \mathrm{d}x\mathrm{d}y \\
&\quad= \int_A \frac{1}{2\pi} \cdot \mathrm{e}^{-\frac{1}{2}r^2} r \mathrm{d}r\mathrm{d}\theta \\
&\quad= \int_A {\color{red} p_\Theta(\theta) p_R(r)} \mathrm{d}r\mathrm{d}\theta
\end{aligned}$$

$$\left(p_\Theta(\theta) := \frac{1}{2\pi}\mathbb{1}_{\{\theta\in[0,2\pi)\}},\ p_R(r) := r\mathrm{e}^{-\frac{1}{2}r^2}\right).$$

R と Θ は独立で $\Theta \sim U(0, 2\pi)$.

R の分布

R の密度関数は

$$p_R(r) := r\mathrm{e}^{-\frac{1}{2}r^2} \qquad \text{for } r \geq 0.$$

累積分布関数は

$$\begin{aligned} F_R(x) &= \int_0^x r\mathrm{e}^{-\frac{1}{2}r^2}\,\mathrm{d}r \\ &= \left[-\mathrm{e}^{-\frac{1}{2}r^2}\right]_0^x \\ &= 1 - \mathrm{e}^{-\frac{1}{2}x^2} \end{aligned}$$

よって

$$F_R^{-1}(u) = \sqrt{-2\log(1-u)}$$

逆関数法より $\sqrt{-2\log(1-U)}$ は R と同じ分布。U と $1-U$ は同じ分布なので $\sqrt{-2\log(U)}$ は R と同じ分布。

ボックス–ミューラー法

コンピュータ上でのシミュレーションなどで正規分布に従う乱数はよく用いられる。

コンピュータで標準正規分布に従って実数をランダムに発生させたい。

$U(0,1)$ に従ったランダムな実数は発生できると仮定する。

ボックス–ミューラー法

1. $u, v \leftarrow \texttt{gen\_uniform}(0,1)$.
2. $r \leftarrow \sqrt{-2\log u}, \quad \theta \leftarrow 2\pi v$.
3. $x \leftarrow r\cos\theta, \quad y \leftarrow r\sin\theta$.

この手続きで生成された x と y は独立な標準正規分布に従う。

これをくり返すことで独立な標準正規分布に従う乱数をいくらでも作れる。$X \sim N(0,I)$ のとき、$AX + \mu \sim N(\mu, AA^T)$ なので一般の多変量正規分布に従う乱数も作れる。例えば $A = \sqrt{\Sigma}$ とすれば $N(\mu,\Sigma)$ に従う乱数が発生できる。

直感的な理解と一般化

二次元標準正規分布の密度関数は $\sqrt{x^2+y^2}$ だけで決まる。よって、$\begin{bmatrix} x \\ y \end{bmatrix}$ を回転しても密度関数の値は不変。一般に $\begin{bmatrix} X \\ Y \end{bmatrix}$ の密度関数 $p(x,y)$ が $r(x,y)=\sqrt{x^2+y^2}$ だけで決まるとき、$p(x,y)$ を $q(r)$ と書くことにすると

$$
\begin{aligned}
\Pr(R \in A,\, \Theta \in B) &= \int_{r(x,y)\in A} \int_{\theta(x,y)\in B} p(x,\, y)\mathrm{d}x\mathrm{d}y \\
&= \int_{r\in A} \int_{\theta\in B} q(r) r \mathrm{d}r \mathrm{d}\theta \\
&= \left(\int_{\theta\in B} \frac{1}{2\pi} \mathrm{d}\theta \right) \left(\int_{r\in A} 2\pi r q(r) \mathrm{d}r \right)
\end{aligned}
$$

となる。つまり R と Θ は独立で Θ は $[0, 2\pi)$ 上の一様分布となる。

課題

- 二次元確率変数 $\mathbf{Z} = \begin{bmatrix} X \\ Y \end{bmatrix}$ が正規分布 $N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sqrt{2} & -1 \\ -1 & \sqrt{2} \end{bmatrix}\right)$ に従うとする。このとき $X + 2Y + 1$ が従う分布をもとめよ。

- 二次元確率変数 $\mathbf{Z} = \begin{bmatrix} X \\ Y \end{bmatrix}$ が標準正規分布 $N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ に従うとする。対称行列 $A \in \mathbb{R}^{2\times 2}$ について $\mathbf{Z}^T A \mathbf{Z}$ のキュムラント母関数をもとめよ。A のスペクトル分解を $A = V^T D V$ とするとよい。

- 確率変数 $Z_1, Z_2, \ldots, Z_n$ が独立でそれぞれが標準正規分布 $N(0,1)$ に従うとする。このとき、$\sum_{k=1}^{n} Z_k^2$ のキュムラント母関数をもとめよ。